CHAPTER 31

Peptide Purification by Reverse-Phase HPLC

Chris Shaw

1. Introduction

The increasing sophistication and sensitivity of microsequencing of peptides and proteins in recent years have presented the biochemist with the feasibility of primary structural determinations on low-picomole quantities of material. This has led to a plethora of structural data in many different areas of interest, which would until now have taken decades of effort to produce. Complementary to the increased sensitivity of analytical technology has been the application of HPLC in the rapid and highly discriminating separation of biomolecules. If such fractionations are interfaced with a specific rapid detection system, such as an immunoassay for the peptide or protein of interest, then purification to homogeneity from a crude extract can be accomplished in a matter of days. Purification to homogeneity is an absolute prerequisite for successful and meaningful structural analyses. Reverse-phase HPLC is usually the technique of choice for isolation procedures, but if complex biological material represents the starting point of an isolation, it is wise to perform a degree of sample clean up and/or concentration initially. Much of what applies to peptide isolation from complex biological material is also relevant to purification of synthetic peptides from resin eluates. In the case of purification of a peptide from a complex biological material, such as the tissue of production, an appropriate quantity of such starting material is essential. Initial pilot studies are required to determine this quantity

From: *Methods in Molecular Biology, Vol. 32: Basic Protein and Peptide Protocols*
Edited by J. M. Walker

by employing either immunometric or biometric analysis. Such studies should also determine the extraction medium that produces maximal solubilization and stability of the peptide of interest. Once these factors have been determined, a preparative extraction can be performed. Often a large volume of turbid extracts results, which would be totally unsuitable for HPLC fractionation. Clarification can often be achieved by high-speed centrifugation, and if the extraction medium is organic-solvent-based, concentration can be affected in part by rotary evaporation. Peptides in the remaining solution can be further concentrated by pumping through disposable cartridges of reverse-phase stationary phases arranged in series. Following washing, bound peptides can be eluted with an organic solvent, such as acetonitrile, and the eluate evaporated to near dryness. The peptides that were once contained in perhaps several liters of extraction medium are now contained in a volume of several milliliters. The total peptide concentration in the extract, however, is now very high, remaining unsuitable for HPLC. At this stage, a high-capacity biomolecule separation system such as soft gel permeation chromatography is often employed. This serves several purposes. By choosing an appropriate gel for the peptide of interest, a high degree of purification can be achieved by removal of higher and lower mol-wt contaminants. This fractionation will yield an estimation of molecular mass and may also yield data on peptide heterogeneity. If the gel is eluted with an acid mobile phase, such as acetic acid, fractions containing the peptide of interest can be directly pumped onto a semipreparative HPLC column. Silica is the most commonly employed reverse-phase HPLC support, and this can be modified in many ways with respect to particle size and shape, pore size, and surface chemistry, which is most often C-18, C-8, C-4, or diphenyl. Combining some or all of these attributes in a sequential series of fractionations usually results in purification of the peptide of interest to apparent homogeneity. Column eluates can be monitored simultaneously at various wavelengths. In initial fractionations, fractions can be collected automatically, but as purification proceeds, it is best to collect peaks of absorbance manually. This achieves even higher resolution of peptides with similar retention times. At each stage, fractions or peaks of absorbing material are screened for the peptide of interest until the detection proce-

dure indicates that the material of interest corresponds to a discrete symmetrical peak of absorbance. If the peptide contains tyrosyl or tryptophanyl residues, which absorb strongly at 280 nm, then symmetrical peaks of absorbance at 214 and 280 nm are indicative of a high degree of purity. Structural analysis of an isolated peptide will usually involve an accurate mass determination by some method of mass spectroscopy. This can usually be achieved for 10–50 pmol quantities of peptide. Amino acid composition may be obtained by conventional analysis following acid hydrolysis and derivatization usually with phenylisothiocyanate (PITC) or with phenylthiohydantoin (PTH). These derivatizing reagents react with amino groups and render all amino acids visible in the UV range. Larger quantities of peptide (500–1000 pmol) are usually required for this analysis. Finally, the primary structure of the isolated peptide can be established by automated Edman degradation using a gas-phase sequencer. Sensitivities of such instruments are in the subpicomole range, but usually 200–500 pmol of a peptide are required to establish the full primary structure unequivocally. The quantity of peptide required is often dependent on the mol mass with larger quantities required for longer peptides. Even so, some peptides will require enzymatic fragmentation and sequencing of fragments to determine the full structure. This can be achieved by using one or more highly site-specific endoproteinases, the choice of which is based on preliminary compositional and/or sequence information. By employing such enzymes at 10–50-fold lower molar concentrations than peptide, there is little interference in subsequent structural analyses from amino acids or oligopeptides resulting from enzyme autodigestion. Usually, peptide digests would be repurified by reverse-phase HPLC, and resulting fragments would be separately subjected to structural analyses.

2. Materials

1. All water employed should be of HPLC grade, either commercially available or produced from a purpose-designed HPLC water system. Periodic checks should be made on the quality of all reagents employed by running blank gradients on chromatographic columns to assess for spurious peaks owing to contaminants. (N.B. Some spurious reagent peaks, if different in retention time from peptides of interest, can act as useful internal standards.)

2. Acetonitrile should be of far UV grade or, better still, the even more UV transparent "super gradient" grade as offered by some commercial companies. All HPLC solvents should be filtered and degassed with helium either before daily usage or, as facilitated by some equipment, during chromatography.
3. Several different counterions can be employed, and these should be used in the same concentration (usually 0.1%, v/v) in both aqueous and organic mobile phases. Trifluoroacetic acid, heptafluorobutyric acid, and phosphoric acid are most commonly employed. For peptides of more acidic character, triethylamine can be employed, but must be buffered to a pH consistent with chromatography on silica-based stationary phases (*see* Note 1). All counterions must be of HPLC grade.
4. Avoid the use of glass or polystyrene tubes during chromatography, since these possess charged groups on their surfaces that can result in a high degree of peptide adsorption. This becomes especially evident as the peptide becomes purer. Use uncharged plastic tubes, such as polyethylene or, ideally, polypropylene tubes, throughout.
5. For gel permeation chromatography, a gel with a suitable fractionation range should be chosen. Sephadex G-50 or the Sephacryl equivalent has such a suitable range for most peptides in the range of 1–30 kDa. The elution buffer may vary, but it is important not to add protein since this defeats the purpose of peptide purification. As a consequence, chosen eluants should facilitate high peptide recoveries, and for this reason acetic acid, in the range of 0.5–2*M*; has been used extensively with a high rate of success. The higher molarity solutions inhibit bacterial growth, and once the column has been calibrated with suitable markers, facilitate numerous fractionations.
6. HPLC equipment should include an elution gradient programmer, a solvent delivery system, two detectors, one of which monitors column effluents at 280 nm for detecting peptides containing tyrosine and tryptophan and the other at 214 nm for detecting all peptides. Column effluents should be collected using a programmable fraction collector.
7. HPLC columns can contain C-4, C-8, C-18, or diphenyl-derivatized stationary phases for reverse-phase fractionations. In initial stages, semipreparative columns can be employed (1 × 60 cm or 1 × 30 cm) followed by analytical columns (0.46 × 25 cm) as the peptide becomes purer. For ion-exchange HPLC fractionations, either cation- or anion-exchange columns can be employed, but organic modifiers, such as acetonitrile should be added to both start and elution solvents in the range of 20–30% (v/v).

8. Endoproteinases employed for peptides of more than 30 amino acid residues, which are difficult to sequence directly, should be of sequencing grade. Manufacturer's instructions, supplied with each batch, should be adhered to.

3. Methods

3.1. Preparation of Crude Extract

1. Once suitable pilot experiments have been performed to determine the appropriate extraction medium for a given peptide and the quantity of starting material required, a preparative extraction for peptide purification should be performed. Tissues should be homogenized in extraction medium, and maintaining this at a low temperature (4°C) will aid the inhibition of endogenous peptidase activity. For many peptides, ethanol/0.7*M* HCl (3:1, v/v) is a highly efficient extraction medium, and this should be employed at a ratio of 8 vol/g of tissue or 3 vol/mL of plasma, serum, culture medium, or other biological fluid. Homogenates or other biological fluid extracts should be constantly stirred and kept at 4°C for 12–18 h to ensure efficient solubilization of peptides.
2. The volume of the extract should be reduced by organic solvent evaporation and/or lyophilization where appropriate, and peptides in the remaining solution should be concentrated by pumping through disposable C-18 cartridges arranged in series. Low flow rates of 10–12 mL/h ensure a high degree of adsorption. After washing in 0.1% (v/v) aqueous trifluoroacetic acid, bound peptides can be eluted from the columns with acetonitrile. Eluants can then be evaporated to *near* dryness (*see* Note 2).
3. Peptides generated by chemical or enzymatic cleavage of purified proteins (Chapters 32 and 33) can be subjected to chromatographic fractionation directly.

3.2. Gel Permeation Chromatography

1. Gel permeation chromatography, with a 2*M* acetic acid mobile phase, can be carried out successfully at room temperature with no significant losses in peptides.
2. Prior to loading lyophilized samples onto these columns, reconstitution in mobile phase (2–3 mL) followed by a short high-speed centrifugation step to remove microparticulate matter, is advised since this latter step greatly increases the working life of the column. A 90 × 1.6 cm column should be eluted at a flow rate of 10–12 mL/h to facilitate component partition.

3. Fractions should be collected at 15-min intervals and, after the predetermined total column volume has been collected, a small aliquot of each should be subjected to a suitable detection system, such as bioassay or radioimmunoassay, for the peptide of interest.

3.3. Reverse-Phase HPLC

1. Gel permeation chromatographic fractions containing the peptide of interest should be pooled and, if in an acetic acid mobile phase, can be pumped directly onto a semipreparative reverse-phase HPLC column that has been equilibrated in starting solvent, such as 0.1% (v/v) aqueous TFA. Generally, the size of column chosen at this stage should relate to the mass of starting tissue employed in the extraction, since this reflects the peptide loading, which in turn will affect resolution. Ideally, the column should be at least 1 × 30 cm, preferably 1 × 60 cm, but employment may be affected by the capital cost. For peptides >4 kDa in mol mass, C-18 chemistry is appropriate at this initial stage, because this then permits the subsequent employment of wide-pore (300 Å) lower carbon-loaded analytical columns. These greatly improve the resolution of peptides of this size. The flow rates employed in this semipreparative stage should be appropriate for column dimensions and are usually in the range of 2–4 mL/min. Gradients are usually linear running from TFA/water (0.1:99.9, v/v) to TFA/water/acetonitrile (0.1:29.9:70.0, v/v/v) in 70 min.
2. To obtain a peptide of apparent homogeneity, sequential chromatography on wide-pore C-8, C-4, or diphenyl analytical columns is required. At each stage, fractions containing the peptide of interest should be pooled and diluted 1:4 with initial aqueous mobile phase to remove the eluting potential of acetonitrile.
3. After pumping of this diluted peptide solution onto the column, one should wait until the absorbance returns to baseline (usually 15–20 min) before initiating the elution gradient. Gradients employed for elution will depend on the hydrophobicity of the peptide of interest. Generally, the gradient should rapidly progress to approx 15% of eluting solvent less than that required for peptide elution and then proceed at approx 0.5% or less of eluting solvent/min. These shallow gradients permit resolution of peptides with similar hydrophobic properties.
4. On analytical columns, low flow rates of 1 mL/min or less favor partition of peptides with similar hydrophobic properties.
5. On all analytical runs, column effluents should be monitored at several wavelengths. Although diode array detection would be the ideal, much

information about peptide composition and degree of purity can be obtained by simultaneous monitoring at 214 and 280 nm either by means of a dual-wavelength detector or two fixed-wavelength detectors arranged in series. Although the 214-nm-detector will detect all peptidic material, the 280 nm detector will indicate the degree of aromaticity of the peptide. With the 280-nm detector set at a sensitivity 10 times higher than the 214-nm detector, a single tyrosyl side chain yields a similar deflection to approximately five peptide bonds. A single tryptophanyl side chain on a peptide yields approximately three times this absorbance. Toward the end of a purification scheme, peaks of peptide absorbance should be collected manually. This ensures that closely eluting contaminants are not collected into the same tube as the peptide of interest, which may occur if automatic fraction collection is employed. The delay between detection and elution from the end of the chromatography tubing can be ascertained in any fixed system by calibration with a known peptide standard, which can be detected in manual fractions either immunochemically or by bioassay.

6. For peptides that are difficult to purify to homogeneity by this standard scheme (i.e., using trifluoroacetic acid as counterion), different counterions, such as phosphoric acid or heptafluorobutyric acid, may be employed (*see* Note 3). These counterions render the peptide more hydrophilic and more hydrophobic, respectively, but also alter the selectivity of peptide interaction with the stationary phase often producing baseline resolution of peptides that cannot be separated conventionally. Shaw et al. *(1)* illustrates the resolution obtained by changing counterions. Ion-exchange HPLC may also be employed in troublesome cases, often with remarkable resolving power (*see* Note 4). Leung et al. *(2)* describe an application of cation-exchange HPLC to solve a problem of peptide purification. Many such columns, however, although excellent for protein chromatography, bind small peptides by hydrophobic interaction in small hydrophobic pockets. These may be difficult to recover. Incorporation of 20–30% acetonitrile in both starting and eluting mobile phases often overcomes this effect. However, if salt or buffer concentration gradients are employed, the concentration in the eluting solvent will be reduced by the presence of organic modifier. This may mean that the final concentration employed may not be of sufficient strength to elute the peptide of interest. If the peptide is eluted from this system, it is usually highly pure if the ion-exchange run has been carried out toward the end of a reverse-phase purification scheme. Ideally, a final reverse-phase fractionation should be carried out to facilitate removal of ion-exchange buffer salts.

7. If peptides are forwarded to a core facility for structural analysis, the samples should be sent in a sealed polypropylene tube in the elution solvent and should not be lyophilized. This may result in significant losses. When received for structural analysis, the sample can be evaporated to a volume appropriate for mass spectroscopy or gas-phase sequencing by direct application. The peptide will thus not have been subjected to lyophilization at any stage of the purification procedure, except when present initially in the crude extract.

3.4. Endoproteinase Digestion

1. In some cases, where the peptide of interest is above 30 amino acid residues in size or of unusual structure, direct gas-phase sequencing may not, because of sequential sequencer losses, result in elucidation of the entire primary structure. Mass spectroscopy data will enable estimation of the approximate size of the segment not sequenced by computation of the mass of the primary structure deduced by the gas-phase sequencer. The primary structural information obtained will enable the choice of a suitable endoproteinase to be made. A range of such endoproteinases is available commercially in highly purified sequencing grades, and the choice will depend on aspects of the primary structure of each individual peptide. Trypsin, chymotrypsin, endoproteinases Asp-N, Glu-C, Lys-C, and Arg-C are a few of those available (*see* Chapter 32).
2. After incubation with a suitable specific endoproteinase, the digest is fractionated by reverse-phase HPLC, and peptide fragments are collected manually. Each can then be subjected to structural analyses, when the full primary structure of the peptide of interest can be determined (*see* Note 5). McKay et al. *(3)* and Maule et al. *(4)* describe the use of such endoproteinases to elucidate the full primary structure of peptides.
3. In other circumstances, when there is a high degree of probability that an isolated peptide has an N-terminal pyroglutamate group (caused by cyclization of an N-terminal glutamine in the presence of acid, either within secretory granules or in the isolation procedure), incubation of the peptide with pyroglutamate aminopeptidase is a necessary prerequisite to automated Edman degradation, which will not work if the peptide has this blocked N-terminus. The deblocking reaction proceeds rapidly (1 h at 37°C), after which the digest can be directly injected onto a reverse-phase HPLC column. If the same gradient as used in the final purification is employed, then the elution of the peptide a few

minutes earlier is indicative of successful deprotection. Gas-phase sequencing can then be performed. The primary structures of possum *(1)* and frog *(5)* neurotensins were deduced using this methodology.

4. Notes

1. The matrix of silica-based HPLC columns, as indicated in manufacturer's operating instructions, is damaged by mobile phases whose pH is around or above neutral. When employing basic counterions, such as triethylamine, ensure that mobile-phase pH values are below 7. If pH values above this are required, then reverse-phase columns with an organic polymer matrix can be utilized.
2. When lyophilizing extracts, it is important not to achieve complete dryness since this will often produce a protein/peptide pellet that is difficult to redissolve resulting in significant peptide losses.
3. Because of the suppression of 214-nm absorbance of peptides when heptafluorobutyric acid is employed as a counterion, this should be employed during the early stages of fractionation when absorbances are not of importance. Other counterions, such as trifluoroacetic acid or phosphoric acid, can be employed in later stages. However, it is not recommended to employ phosphoric acid at the final stage, especially if the peptide sample is to be lyophilized prior to structural analysis. Phosphoric acid is not volatile, and the resultant high concentration of acid at the latter stages of drying may hydrolyze the sample.
4. Extreme caution should be exercised when employing ion-exchange HPLC columns. Such columns are normally stored in a high concentration of organic solvent (20–50%) containing a weak solution of appropriate counterion. Always ensure that columns are washed clean of high molarity elution buffers prior to storage, or precipitation in the pumps and mixing chambers may occur. This may also occur on the column and, if the column is in a glass package, explosive shattering may occur.
5. The use of endoproteinases and the methodology for isolation of peptide fragments described in this section can also be employed for primary structural analysis of large proteins. Fragments of large proteins may also be generated by chemical means, such as by cyanogen bromide. For endoproteinase digestion, sequencing grades of appropriate enzymes should be employed to ensure against the risk of aberrant cleavages.
6. A schematic representation of a hypothetical novel neuropeptide isolation and characterization is shown in Figs. 1–4.

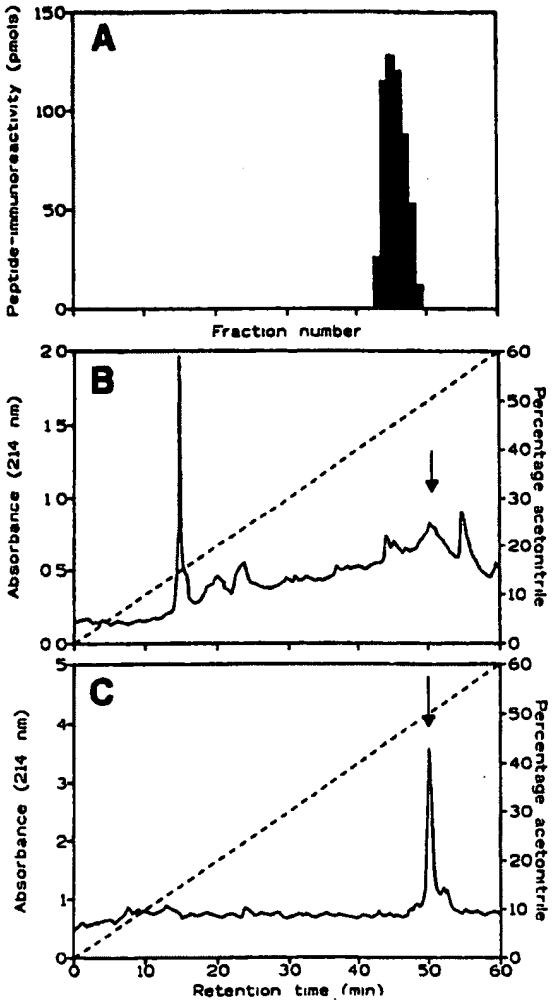

Fig. 1. (A) Gel permeation chromatogram (Sephadex G-50 [fine]; 90 × 1.6 cm; mobile-phase 2*M* acetic acid; flow rate 10 mL/h) of a crude tissue extract containing a novel neuropeptide. Fractions containing the neuropeptide were localized by radioimmunoassay of a small aliquot of each fraction. The antiserum employed for radioimmunoassay was raised to a known peptide and crossreacted with the novel neuropeptide. (B) Semipreparative (C-18) reverse-phase HPLC chromatogram (Whatman Partisil ODS3; 60 × 1 cm) of gel permeation chromatographic fractions containing neuropeptide immunoreactivity. Aliquots of fractions were screened by radioimmunoassay as before, and the retention time of the novel neuropeptide is indicated by an arrow. (C) Analytical (C-8) reverse-phase HPLC chromatogram (Vydac 208TP54; 25 × 0.46 cm) of the novel neuropeptide resolved by semi-preparative chromatography Immunoreactivity is indicated by an arrow. HPLC elution gradients are indicated by hatched lines, and the starting solvent in each case was 0 1% (v/v) aqueous TFA.

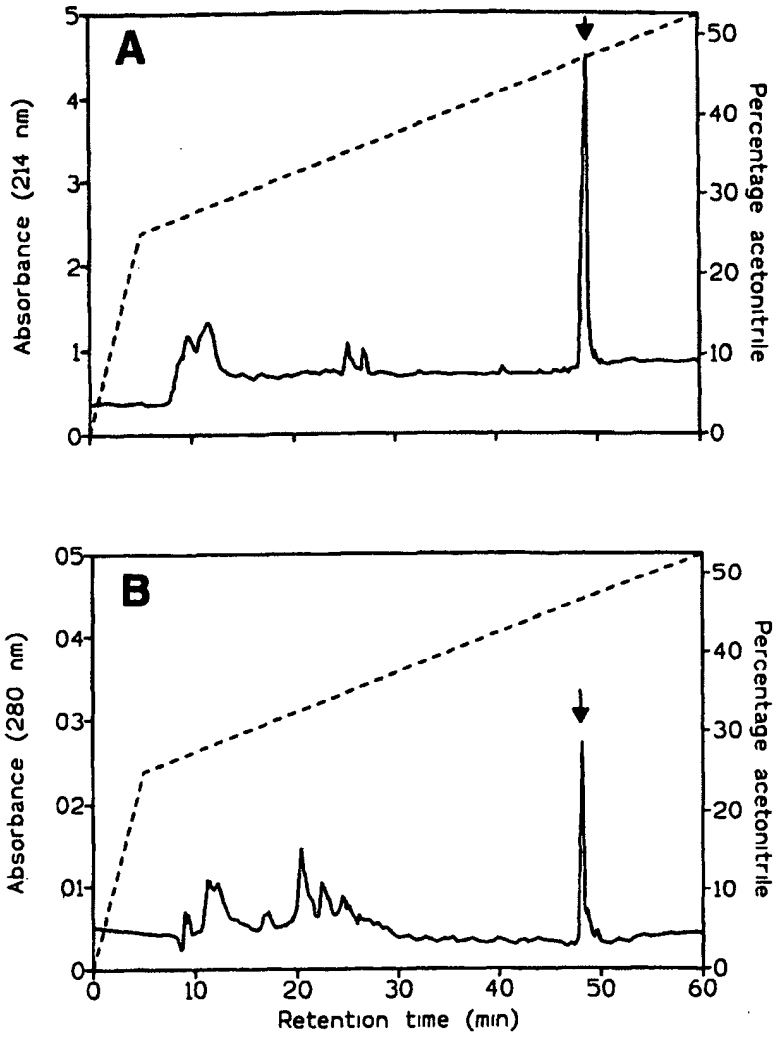

Fig 2. Analytical (C-4) reverse-phase HPLC chromatogram (Vydac 214TP54; 25 x 0.46 cm) of the final fractionation of the novel neuropeptide. The column effluent was monitored at 214 nm (A) and 280 nm (B). Note the 10-fold difference in sensitivity of both wavelengths. Peaks of absorbing material were collected manually, and immunoreactivity corresponded to the major peak of absorbance, which is marked by an arrow. The 280-nm absorbance of the peptide indicates the presence of aromatic amino acid residues, and the symmetry of absorbance peaks at both wavelengths indicates a high degree of purity The elution gradient is indicated by the hatched line, and the starting solvent was 0.1% (v/v) aqueous TFA

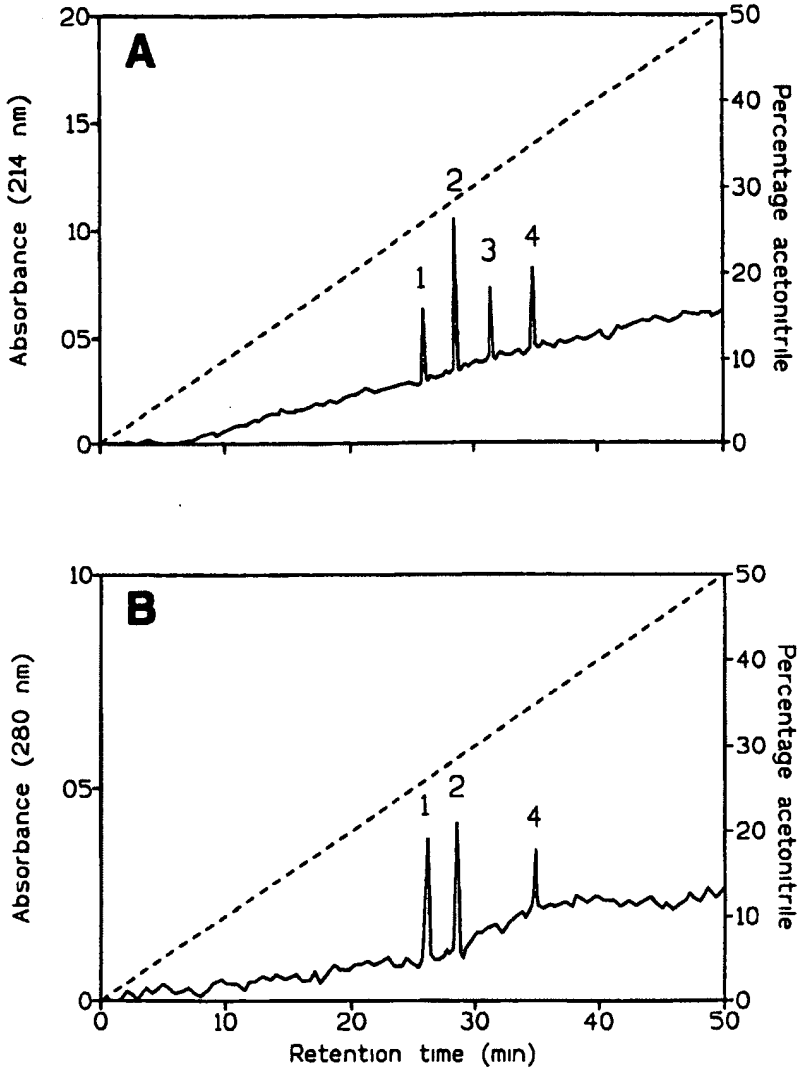

Fig 3. Analytical (C-8) reverse-phase HPLC chromatogram (Vydac 208TP54; 25 x 0.46 cm) of an endoproteinase Glu-C digest of the novel neuropeptide purified to apparent homogeneity as illustrated in the previous figure. The absorbance of the column effluent was monitored at 214 nm (A) and 280 nm (B) Four fragments (F1–F4) of the novel neuropeptide were baseline-resolved, and note that F3 has no absorbance at 280 nm, indicating that it does not contain aromatic amino acid residues The structural characterization of these fragments is schematically represented in Fig. 4. The elution gradient is indicated by a hatched line, and the starting solvent was 0 1% (v/v) aqueous TFA.

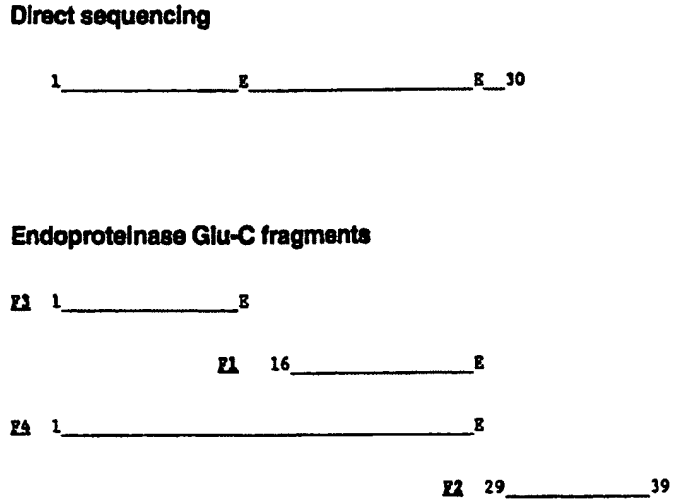

Fig. 4. Residues 1–30 of a novel neuropeptide were assigned following a single direct gas-phase sequencing run The computed molecular mass of this peptide was less than that derived by mass spectroscopy, indicating that 9–12 amino acid residues at the C-terminal end of the peptide had not been obtained. Two glutamic acid residues (E in single-letter notation) were present at positions 15 and 29, respectively. A second batch of peptide was digested with endoproteinase Glu-C, and the profile of peptide fragments obtained is shown in Fig. 3. Sequencing of the peptide fragments F1–F4, as indicated schematically, resulted in the deduction of the full primary structure of a 39 amino acid residue peptide whose computed mol mass was in agreement with that obtained by mass spectroscopy.

References

1. Shaw, C., Murphy, R., Thim, L , Furness, J B , and Buchanan, K. D. (1991) Marsupial possum neurotensin: a unique mammalian regulatory peptide exhibiting structural homology to the avian analogue. *Regul. Pept.* **35,** 49–57.
2. Leung, P. S., Shaw, C., Maule, A. G., Thim, L., Johnston, C. F., and Irvine, G. B. (1992) The primary structure of neuropeptide F (NPF) from the garden snail, *Helix aspersa. Regul. Pept.* **41,** 71–81.
3. McKay, D. M., Shaw, C., Thim, L , Johnston, C. F., Halton, D. W., Fairweather, I., and Buchanan, K. D. (1990) The complete primary structure of pancreatic polypeptide from the European common frog, *Rana temporaria. Regul. Pept.* **31,** 187–198.
4. Maule, A. G., Shaw, C., Halton, D. W., Thim, L., Johnston, C. F., Fairweather, I., and Buchanan, K. D (1991) Neuropeptide F: a novel parasitic flatworm regulatory peptide from *Moniezia expansa* (Cestoda: Cyclophyllidea) *Parasitology* **102,** 309–316.
5. Shaw, C., McKay, D M., Halton, D. W., Thim, L , and Buchanan, K D (1992) Isolation and primary structure of an amphibian neurotensin. *Regul. Pept* **38,** 23–31